SPECIAL ISSUE PAPER

AI inspired EEG-based spatial feature selection method using multivariate empirical mode decomposition for emotion classification

Muhammad Adeel Asghar[1] · **Muhammad Jamil Khan**[1] · **Muhammad Rizwan**[2] · **Mohammad Shorfuzzaman**[3] · **Raja Majid Mehmood**[4]

Abstract

Classification of human emotions based on electroencephalography (EEG) is a very popular topic nowadays in the provision of human health care and well-being. Fast and effective emotion recognition can play an important role in understanding a patient's emotions and in monitoring stress levels in real-time. Due to the noisy and non-linear nature of the EEG signal, it is still difficult to understand emotions and can generate large feature vectors. In this article, we have proposed an efficient spatial feature extraction and feature selection method with a short processing time. The raw EEG signal is first divided into a smaller set of eigenmode functions called (IMF) using the empirical model-based decomposition proposed in our work, known as intensive multivariate empirical mode decomposition (iMEMD). The Spatio-temporal analysis is performed with Complex Continuous Wavelet Transform (CCWT) to collect all the information in the time and frequency domains. The multiple model extraction method uses three deep neural networks (DNNs) to extract features and dissect them together to have a combined feature vector. To overcome the computational curse, we propose a method of differential entropy and mutual information, which further reduces feature size by selecting high-quality features and pooling the k-means results to produce less dimensional qualitative feature vectors. The system seems complex, but once the network is trained with this model, real-time application testing and validation with good classification performance is fast. The proposed method for selecting attributes for benchmarking is validated with two publicly available data sets, SEED, and DEAP. This method is less expensive to calculate than more modern sentiment recognition methods, provides real-time sentiment analysis, and offers good classification accuracy.

Keywords Artificial intelligence · EEG-based emotion classification · Feature extraction · Signal processing · Health care

1 Introduction

It is important to monitor mental stress levels to prevent and treat chronic diseases early on. Various techniques have been developed to recognize stress levels. Some methods are based on physiological signals called an electroencephalogram (EEG) [1]. EEG signals are the most widely used method of recognizing human emotions. Because the behavior of the EEG signal is not linear, there are many problems with accurately recognizing emotions [2]. The EEG signals are thus converted into a two-dimensional Spatio-temporal spectrogram. Afterward, analyze

✉ Raja Majid Mehmood
  rmeex07@ieee.org; rajamajid@xmu.edu.my

  Muhammad Adeel Asghar
  adeel.asghar@students.uettaxila.edu.pk

  Muhammad Jamil Khan
  muhammad.jamil@uettaxila.edu.pk

  Muhammad Rizwan
  muhammad.rizwan@uettaxila.edu.pk

  Mohammad Shorfuzzaman
  m.shorf@tu.edu.sa

[1] Telecommunication Engineering Department, University of Engineering and Technology, Taxila, Pakistan

[2] Computer Engineering Department, University of Engineering and Technology, Taxila, Pakistan

[3] Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

[4] Information and Communication Technology Department, School of Electrical and Computer Engineering, Xiamen University Malaysia, Sepang 43900, Malaysia

the effects of emotional behavior using guided learning techniques used in ImageNet [3]. The best way to process non-linear EEG signals in the time and frequency domain. One area of emotion recognition is the use of deep neural networks, which are useful for estimating complex functions. Therefore, recognizing human emotions is a difficult task. Also, different people behave differently in the same emotional state [4]. Emotional states can be derived from external and internal human responses. By influencing the development and computation of advanced machine learning tools, many methods can be used to detect human emotions using brain wave signals. BCI (brain–computer interface) is a mechanism that first collects a signal using a bio-transducer. BCI can play an important role in the provision of health services in smart cities [5]. The signals collected by BCI can help better understand your emotional response. However, it is not clear how to decipher their emotions accurately and broadly.

EEG-based emotion recognition is designed in real-time applications and can provide real-time patient monitoring services using physiological signals. It is a system like [6], a cloud-enabled cyber-physical patient monitoring location system that uses smartphones to collect voice and EEG signals in a scalable, real-time, and efficient way. The coronavirus (COVID-19), a pandemic that recently hit the world, can also affect human emotions. Sentiment analysis also provides a gateway to the study of emotional behavior in patients with corona [7]. The effect of social networks on emotions is also an application of the user's ability to change their emotional state [8]. Due to the great popularity of social networking sites, it is helpful for users to give feedback on many social events. This makes it easy to create, share, and distribute social events, and as a result, users provide large amounts of multimedia data (for example, images, videos, and text) for various actual events in different shapes and sizes.

With regard to edge computing, there is a research tendency to estimate the emotions caused by the visualization of stimuli in different IoT applications. This paper exploits the advantage of deep neural networks to extract features using AlexNet, ResNet-50, and Inception V3 from all channels related to emotions. Due to a large number of input data sources, it takes a long time to train the network. To reduce computational cost and get better results, this article suggested channel selection and a high-quality feature selection algorithm for fast processing. A channel selection algorithm based on differential entropy and mutual information (DEMI) is proposed to select the number of channels with emotional information. The combination of the proposed model differs from the bull technique to choose the most suitable features for the emotion, which helps to reduce the ultimate feature dimension. EEG-based sentiment classification model that uses the DNN model and feature selection

is validated on SJTU SEED and DEAP data sets. Support vector machine and k-Nearest neighbor are used to classify data into three and four dimensions for SEED and DEAP data sets, respectively. The proposed model achieves better accuracy as compared to other advanced methods of human emotion recognition.

Recent research shows that detecting emotions using EEG signals is not an easy task. Many researchers have tried to recognize emotions. We use two commonly used SEED and DEAP data sets for model validation and comparison with recent techniques. Based on these two data sets, we discuss the latest research on emotion recognition. In the literature, the authors [9] were used to classify EEG signals using the radial biased function (RBF) based on the kernel support vector machine (SVM). They reached a precision of 60% with the proposed method. Reference [10] extracts spatial features using a two-dimensional circular space. They used the Dual-Tree Continuous Wave Transform (DT-CWT) to observe the time and frequency behaviour of the EEG signal. They trained the model using a simple recurring unit (SRU) to extract functions. For the best sort of performance, they rate features with k-Nearest Neighbor (K-NN). The CNN method is used to detect EEG-based seizures. The authors of [11] used deep CNN to extract features and confirmed the presence of epileptic seizures. Cloud-based multilevel perception (MLP) [12] also reads emotional behavior for real-time applications in smart cities. EEG signals are a mixture of different frequency bands. This link reflects the specific cognitive state. [12] achieves an accuracy of more than 89% using a convolutional network and an MLP with 2 hidden layers.

Correlation-based subset selection (CSS) helped select the desired EEG signal characteristics. The authors in reference [13] used higher-order statistics to calculate the efficiency of the system. To reduce the dimension of the obtained feature vector, CSS was claimed to achieve superior sort performance compared to conventional PCA (Principal Component Analysis). EMD (Empirical Mode Decomposition)-based decomposition method helps researchers analyze raw EEG signals to reduce dimensional curses. The EMD is divided into a smaller set of functions called IMFs (Intrinsic Mode Functions) [14–16]. The higher the MFI value, the lower the signal distortion. This will help a lot in terms of computational costs. Similarly, [17] demonstrates that splitting CSS with OPA (Ordinal Pattern Analysis) increases sort accuracy by up to 16%.

Entropy-based selection of functions is also a major concern for many researchers. Entropy measures the predictability of features that occur in RAW EEG signals. The spectral entropy of the power is calculated as a feature vector [18]. Eight emotional states were recognized with this entropy function. In [19] also uses EMD-based signal decomposition. Subsequently, the classification performance of the

proposed model was calculated using the matching technique. In [20] and [21] used multivariate EMD (MEMD) and bi-variate (BEMD) to divide the signal into multiple and bi variants. In [22] PSD (power spectral density) and higher-order statistics were also used to extract characteristics of the EMD-based decay signal.

The selection of functions is an important point in the recognition of human emotions. This step is very important to extract high-quality functions for a good sentiment evaluation. The choice of an existing location can be based on consent, mutual information, statistical analysis, and learning savings [23]. These methods are not efficient and you need information before making a decision. In Reference [24] the authors used non-negative Laplace values to estimate the values of the characteristics. Also, the author of [25] used PCA to select the attribute value based on the unique identification. Fisher's Discriminant Index (FDR) [25] is also a traditional feature selection method that selects high-quality features based on high FDR values. Also, [26] proposes a model to optimize the set of using statistical features. The sparse selection algorithm represented by [27] selects attribute properties based on sparse learning methodology. The selected function is used to model sentiment rankings with fair ranking performance for the SEED and DEAP data sets.

Multiscale Information Analysis (MIA) for EEG signals introduces a new function. In [19] uses these characteristics in Russell's circumplex model to evaluate emotional states in four dimensions. The domain's multiple-scale complexity index was its ability to detect emotional states [22, 28]. Based on the previous literature and the sentiment classification method, a rapid emotion recognition method is needed that uses advanced machine learning tools. In this article, we have proposed an EMD-based fast decomposition method to improve the decomposition of the EEG signal. Such decomposition helps to recognize emotions using three pre-trained models in a short period of time with generally good classification performance. Our contribution in EEG-based fast recognition is described below.

1.1 Contributions

Here is a summary of our contributions to this work.

- The Raw EEG signal is further decomposed using EMD-based decomposition known as iMEMD (intensive Multivariate Empirical Mode Decomposition). This is an extended version of MEMD. This reduces the impact of artifacts.
- The degraded EEG signal is transformed into a two-dimensional spectrogram to extract separate spatial features from three pre-trained networks.
- The spatial features of the three networks are defuse together to form a single feature vector. The redundant functions are also eliminated.
- High-quality features are selected with DEMI (Differential Entropy and Mutual Information). Entropy-based channel selection, along with mutual information, is used to select optional channel functions. The output feature vector is reduced to 80%.

The remainder of this paper is as follows: Proposed Framework of multi-model feature selection method is described in Sect. 2. Section 3 describes the experimental setup used to recognize emotions, followed by the experimental results In Sect. 4 achieved using our proposed method. Finally, the work is concluded with the future direction in Sect. 5.

2 Paradigm of the proposed emotion recognition method

Electroencephalography (EEG)-based sentiment analysis requires intensive processing due to the non-linearity of the signal. Hence, large feature vectors created during feature extraction increase the computation cost of the system. To solve this problem, we need a suitable way to select and reduce feature dimensions. In the proposed model, the signal was first broken down into smaller components to avoid scalability. In addition, the features from the three deep neural networks are combined and the redundancy is removed. The combined features are then reduced using clustering methods. Groups of similar functions and resources of groups are considered as a single function. Then we can use this method to get the histogram features. The proposed technique is discussed in the next section. The complete step by step framework of our work is represented in Fig. 1.

Fig. 1 Framework of the proposed work
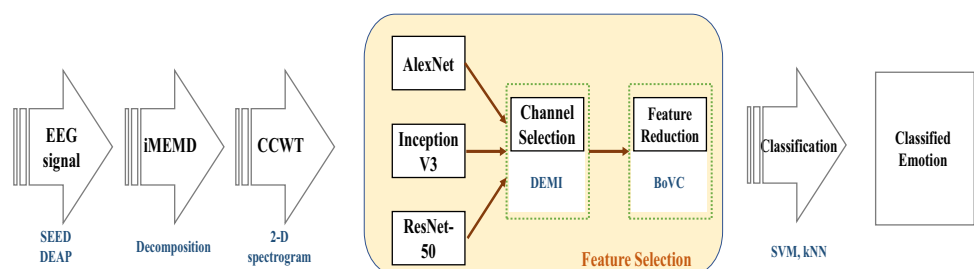
2.1 Raw EEG signal decomposition

Raw EEG signals are so noisy that they have problems acquiring features and recognizing emotional states effectively. Techniques based on empirical mode decomposition (EMD) [20] are used to decompose the signal into various eigenmode functions known as Intrinsic Mode Functions (IMFs). As discussed in the literature work, Bi-variant Empirical Mode Decomposition (BEMD) and Multi-variant Empirical Mode Decomposition (MEMD) are additional types of EMD. In this study, we proposed an expanded version of MEMD known as intensive Multivariate Empirical Mode Decomposition (iMEMD) for rapid decomposition. Unlike MEMD, iMEMD uses an average value for envelope calculation. IMFs are calculated in almost half the time compared to MEMD. The iMEMD decreases the quality of the signal but increases the quality of the function. This method approaches a fast decomposition that takes up half the limit. Rather, this degrades the quality of the output signal, but the characteristics extracted from the method yield much better results than conventional EMD-based techniques. Here is the proposed iMEMD-based decomposition algorithm 1.

The algorithmic steps in iMEMD are somehow similar to EMD and MEMD. Equation 1 is the final equation for calculating iMEMD IMFs. The 5th IMF results in good classification performance that is we use y[5] to calculate our 5th IMF. The middle value of the envelope deteriorates the quality of the output signal. Otherwise it works the same as MEMD, but reducing the envelope size reduces the computational complexity of the system.

$$y[n] = \sum_{i=0}^{N-1} [\epsilon_i(n) + R_\sigma(n)]. \tag{1}$$

2.2 Spatio-temporal analysis

The one-dimensional degraded EEG signal is used to convert it into a two-dimensional image display known as Time–Frequency Representation (TFR). The TFR is collected using CCWT (Complex Continuous Wavelet Transform) [29] for the extraction of spatial features (described in the next section). To acquire a complete knowledge of signal in spatial and temporal domain CCWT is used. Since EEG signals behave non-linearly, the complex continuous wavelet transform is suitable for this type of signal. As described in the literature section, many wavelet-based techniques are used

---

Algorithm 1: intensive Empirical Mode Decomposition Algorithm

1 **Input: Raw EEG signal** $x(t)$
2 **Output: IMFs:** $\sum_{i=0}^{N-1} [\varepsilon_i(n) + R_\sigma(n)]$
  Data: Training and Testing Set for SEED and DEAP
3 $\sum_{i=1}^{\infty} := 0$ `// i= 0,1,2,3....N-1, where N = number of samples`
  `/* Step 1: Choose a sampling set                              */`
  `/* Step 2: Determine a vector projection` $(v(t))$ `w.r.t time for input`
  `    signal` $x(t)$ `                                          */`
  `/* Step 3: Calculate half the Minima` (e_i) `and Maxima` (E_i) `, extremas` $(ext(n))$
  `    */`
4 $ext(n) = \frac{(E_i)}{2}$ && $\frac{(e_i)}{2}$ `/* Step 4: Calculate multivariate envelopes` $env(n)$ `     */`
5 $env(n) = \frac{1}{2} \sum_{i=0}^{N-1} ext(n)$ `/* Step 4: Use the shifting property to describe`
  `    envelopes                                               */`
6 **if** $env(n) == \varepsilon(n)$ **then**
7 | $y(n) = x(n) - env(n)$ `// IMF y(n) created would be`
8 | **if** *Otherwise* **then**
9 | | $y(n) = 0$

10 **else if** $env(n) == R_\sigma(n)$ **then**
  | `/* IMFs without residue                                     */`
11 | **for** *p= 0:N-1* **do**
12 | | $y(n) = \sum_p \frac{|(env_{p-1}(n) - env_p(n))|^2}{(env_{p-1}(n))^2}$

13 **else**
14 | Starts from step 1

15 **while** *IMFs with residue* **do**
16 | $y[n] = \sum_{i=0}^{N-1} [\varepsilon_i(n) + R_\sigma(n)]$

Table 1 Feature vector obtained at each step

	Single network ×4096	Combined network ×4096	RFE ×4096	DE ×4096	MI ×4096	Output FV
SEED	41850	125550	89678	17357	14464	225×36
DEAP	163840	491520	351085	87771	41795	510×48

for time-frequency alignment, but these wavelet-based techniques obtain information related to the sample. CCWT is used in this article for TFR to gain a complete understanding of time limits. Two-dimensional CCWT represents the sparse and characteristic multi-resolution structure of an image. TFR is used to extract features from DNNs. CCWT expresses the signals in terms of wavelet functions which are localized in both time and frequency domain. Continuous EEG signal $x(t)$ for TFR $\chi_\sigma(a, b)$ can be expressed as

$$x(t) = C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi_\omega(a, b) \frac{1}{\sqrt{|a|}} \tilde{\psi} \left(\frac{t-b}{a} \right) ab \frac{da}{a^2}, \qquad (2)$$

where C_ψ is called the wavelet admissible constant whose value satisfies between $0 < C_\psi < \infty$.

2.3 Feature extraction

After complete pretreatment and decomposition, the next step in the proposed method is feature extraction. Features are taken from the 3 pre-trained DNNs listed below. To do this, the two-dimensional TFR images are first scaled to 227×227, then these scaled images are sent separately to 3 DNNs. So we get a table with three feature vectors of these three models. Typical vector sizes obtained from each NN are listed in Table 1. The input of the neural network is the TFR image and output is the feature vector. Features extracted from these DNN are known as spatial features.

2.3.1 AlexNet

AlexNet is a combination of three layers: convolution, maximum pooling and a fully connected layer. The TFR images delivered via a pre-trained AlexNet model. AlexNet has five convolution layer and three fully connected layers. Classification and softmax layer is not used in this work, and we did not perform classification using Alexnet; hence it is only for the extraction of features. AlexNet analyzes each image by pixel value between different emotional states. With a different area of time and frequency in each image, that's why they are more sensitive to various spatial information. Finally,

each function is obtained from a fully connected layer. There are a total of 4096 accommodations using AlexNet. fC7 layer displays the feature vector of each channel.

2.3.2 Resnet-50

The second network used is Resnet-50. As its name implies, it is a residual network with a depth of 50 layers [30]. The model is trained in the same way as AlexNet, but the number of layers and the structure of the neurons is different. Therefore, we collect more effective feature vectors. The network is very useful for extracting non-linear signals due to dropping and hopping. The layers have been stacked and deep convolutional neural networks have changed the classification elements of the image. The scaled TFR images of both data sets are provided separately to the network.

2.3.3 Inception V3

Inception v3 is a widely used image recognition model that has been shown to achieve 78.1% or better accuracy on the ImageNet data set [30]. The model itself consists of symmetric and asymmetric building blocks with layers of convolution, average grouping, maximum grouping, concatenation, failure, and fully connected. This is the last model used for feature extraction. Like other models, it extracts functions up to FC7 layers with 4096 attributes.

2.4 Feature defusing and redundant feature elimination (RFE)

This section describes the decomposition of attributes obtained from the three DNN models. Feature vectors obtained from each neural network are 4185×4096 for data set 1 and 163840×4096 for second data set. Because the classification layer is not used for all three networks, this functionality is added to the fully connected layer 7 (FC7). All three networks have the same FC7 properties, so it is easy to collect the same number of properties from the three networks used. In combination, the defused feature vector is 125550×4096 and 491520×4096 for data sets 1 and 2, respectively. Defused feature vectors can contain similar or redundant features extracted from different models; to reduce this, we further reduced the dimension of the feature vectors
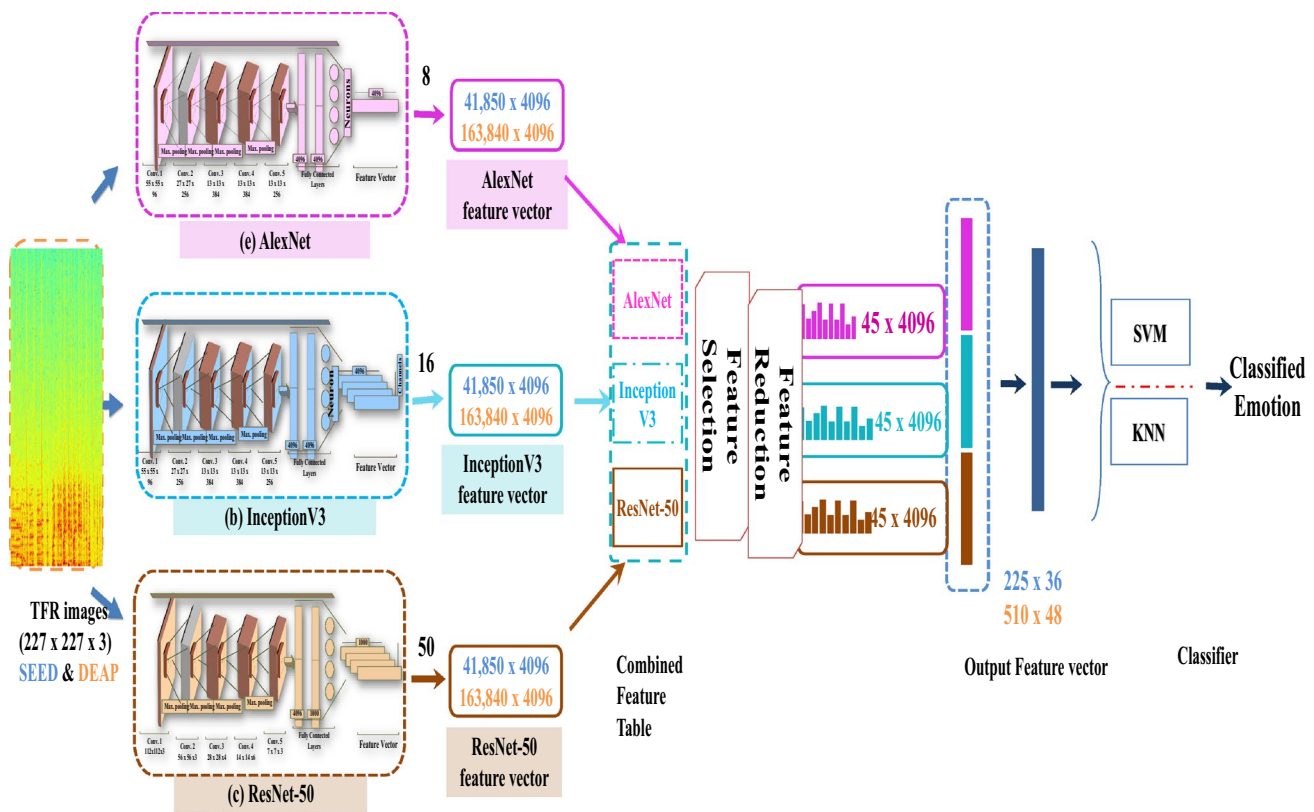
Fig. 2 Multi-model network feature selection

using redundant feature elimination (RFE). This step is very important concerning the computational costs and learning opportunities. The redundancy function not only hinders the classification performance of the proposed model, but also the computation costs. Reducing the functional dimension has also been proven to increase the efficiency of the system. Features are removed based on values using the difference equation. After RFE, the feature vectors are 89678 × 4096 and 351085 × 4096.

2.5 Feature selection

The high-quality features obtained from the three vectors were first compared and redundant features were removed. This is done using the Euclidean distance formula. This also happens simultaneously during treatment. Differential Entropy and Mutual Information (DEMI) is a two-step procedure for selecting additional feature vectors. First, the number of channels is selected using differential entropy, and then the quality characteristic of the selected channel is selected using mutual information of the characteristic values. The two suggested techniques are described below: This reduces the size of the jobs with good sorting performance. Figure 2 shows how to combine multiple networks to select only those functions that have an emotional state. The input

vectors of both data sets are reduced to the highest possible value without disrupting the classification performance of the model.

2.5.1 Differential entropy

Neuroscience explains that a particular region of the brain provides us with different types of waves that lie in bands like alpha, beta, gamma, etc. Pre-frontal, parietal, and temporal areas are where positive valence emotions can be detected [31, 32]. We have used differential entropy to calculate and select the channels being used in our proposed work. Formally, let $(U, C \cup D)$ be an informed decision and $P \subset C$, the differentiation entropy of P with respect to C class is defined by: Formally, $(U, C \cup D)$ is an informed decision and the differential entropy of P for class P *subset* C, class C is defined as:

$$E(P \mid U \oplus C) = -\frac{1}{U} \sum_{x \epsilon U} \log_2 \frac{\left|[x]_C \cap [x]_P\right|}{\left|[x]_P\right|}. \tag{3}$$

A threshold value of 1.725 was set and the channel above that threshold were selected that are; FC1–FC2, FC5–FC6, FP1–FP2, F7–F8, C3–C4. A total of 12 channels were
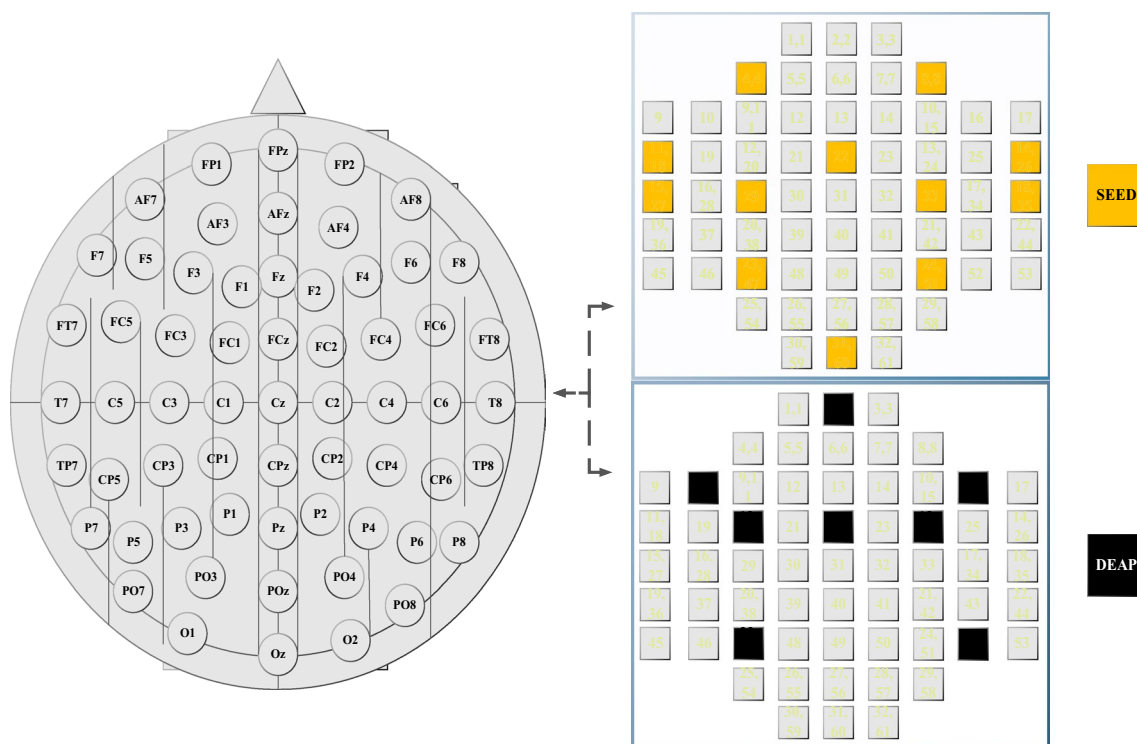
Fig. 3 Selected channels with good classification performance

Fig. 4 Selection of "*k*" based on sum of the squared errors
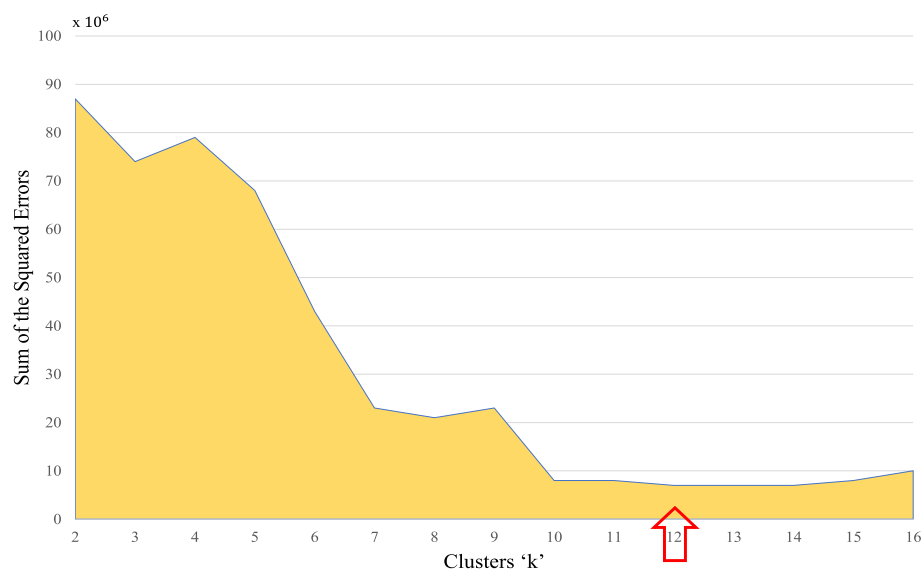
selected out of 62 in case of SEED data set and 8 out of 32 for DEAP data set. Selected channels are shown in Fig. 3.

2.5.2 Mutual information

Mutual information is a way of selecting correlated characteristics. This method works on a selected number of channels. That is, we choose the number of channels to reduce the size of the vector vertically and we use mutual information to reduce the size of the vector horizontally. Select features that correlate between adjacent channels. Equation 4 calculates the probability density function (PDF) of the feature vector and selects the features with correlation. Where *X* and *Y* are the two attribute values.
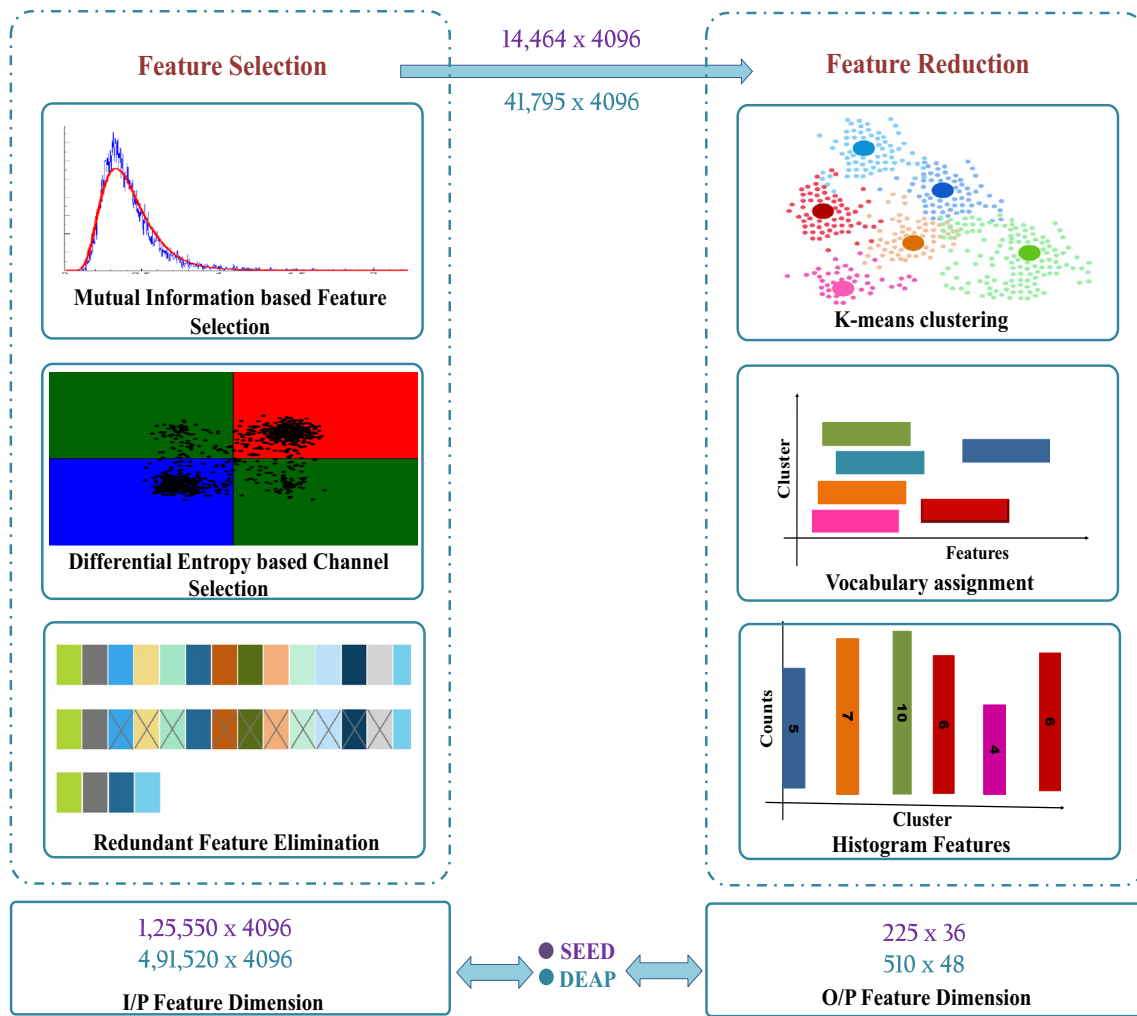
Fig. 5 Two-step feature selection and feature reduction method

$$MI(X;Y) = \sum_{x \in c}^{C} \sum_{y \in c}^{C} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right). \qquad (4)$$

2.6 Feature reduction

The characteristics obtained so far are dimensional, resulting in a dimensional curse in the proposed model. We use k-means to cluster to group similar features. This is an iterative process, the process starts with a random centroid value 'k' and then the average of each centroid is calculated to correctly cluster the features. The goal of k-mean algorithms for different clustering techniques such as hierarchical clustering, fuzzy clustering, density-based clustering, etc., is that these models often suffer from over-fitting in large data sets. For k-mean clustering, it is an iterative clustering algorithm, so it can be easily used for large clustering data sets. The sum of the mean square errors is also calculated to find the

exact value of k after a few iterations. In the graph, Fig. 4, we can see that at $k = 12$ all values are clustered with fewer errors.

2.6.1 Vocabulary assignment

Similar characteristics are clustered by clustering k-means. After correct grouping, all functions were sorted by class, this arrangement is called vocabulary assignment. That is, the traits or emotional states of each class are grouped so that the mean represents an emotional state. This mapping method helps to understand the difference of attributes between different classes, reducing the dimension of the feature vectors (Fig. 5)).

2.6.2 Feature histogram

The next step in reducing the feature dimension is to extract the histogram of the assigned vocabulary. The histogram

feature vector calculates the occurrence of a specific vocabulary feature from the original feature vector. This method automatically removes improperly grouped features, which improves classification performance. The SEED data set contains 3 classes, so the number of columns is $3 \times k = 3 \times 12 = 36$ and for DEAP $4 \times k = 4 \times 12 = 48$. So after reducing the size, the output vectors are 255×36 and 510×48 for the SEED and DEAP data sets, respectively.

3 Experimental setup

In this task, two publicly available data sets are used for validation. The provider collects the data sets in a clean and suitable environment. A digital filter of 0–75 Hz is applied to eliminate the effect of noise. The data sets differed in terms of the subject, the channels used, and the different classes that report emotional state. The details of both data sets are as follows:

3.1 Data set 1:SEED

The proposed model is first validated on publicly available SEED [33] data set developed in China. The data set contains the physiological signals of 15 Chinese participants (8 male and 7 female). The acquired EEG can be used for the multi-model emotion recognition process. Data set was collected from the participants using a bio enhancer placed on a human skull [4] using 10-20 International System [34]. In SEED data set, three emotional states were captured named as positive for calm, negative for fear/stress and neutral for the normal or relaxed condition. This emotional behavior of participants was collected by showing different video clips. SEED data sets utilized three emotional states Positive (Happiness), Negative (Sadness) and Neutral (Relaxed state).

3.2 Data set 2: DEAP

DEAP [35] is also publicly available data set for emotional behaviour evaluation. 32 EEG channels from 32 participants according to international standards, were captured, while participants watch 40 music video clips. Unlike SEED, DEAP data sets used four classes in valence and arousal domain. Then, they perform Awakening, valence,

dominance and liking with self-assessment manikins (SAM) [36] (Table 2).

Both data sets were used to evaluate the results of the proposed model. Three states from the SEED data set and four sentimental states from the DEAP data set were categorized for benchmarking.

3.3 Network training and validation

The proposed feature selection and extraction method are trained on Matlab R2020a using Adam Optimizer with a batch size of 128. We did this with a GTX 1080 GPU of 8GB RAM on a Windows 10 operating system to handle large feature values with less processing. The two data sets are divided into three categories: training, testing, and validation. The model is first trained with the training set, then tested with the test set, and this model is validated for classification performance using the validation set. This configuration is done separately for both data sets. For testing and validation purposes, we use a tenfold cross-validation method for optimal classification. Training multiple networks with nearly 1.2 million layers at the same time is too big for a single GPU. All the parameters of the three previously trained networks are optimized for this type of feature extraction. After selecting entities with smaller dimensions, the validation time to perform sentiment classification with the new data set has also been shortened. The training time in each step of each neural network is given in Table 5.

4 Experimental results

In this section, we examine the implementation of the proposed feature selection method using the classification results obtained with our method. We have performed many evaluation parameters to validate our results. We also compared the results with the retrained model to verify the validity of the proposed method.

4.1 Performance metrics

The following statistics were performed to evaluate the results. To reduce the over-fitting problem, the data set is generally divided into random parts of the same volume. Subsequently, this method is trained with M pieces and

Table 2 Data set comparison

Data set	Participants	Channels	Experiments	Trials	Class
SEED	15	62	5	3	Positive negative neutral
DEAP	32	32	1	40	High arousal high valence low arousal low valence

Table 3 Performance evaluation of SEED data set using SVM classifier

Subject	Accuracy	Precision	Specificity	Recall	F1-score	p-score
Subject 1	78.8	0.79	0.84	0.83	0.89	0.090
Subject 2	84.4	0.87	0.81	0.82	0.95	0.081
Subject 3	84.8	0.90	0.85	0.79	0.87	0.086
Subject 4	83.1	0.93	0.80	0.83	0.96	0.088
Subject 5	91.2	0.89	0.82	0.90	0.95	0.095
Subject 6	87.5	0.93	0.84	0.86	0.92	0.091
Subject 7	74.4	0.82	0.88	0.86	0.83	0.089
Subject 8	84.6	0.83	0.88	0.84	0.88	0.089
Subject 9	78.4	0.97	0.93	0.83	0.96	0.089
Subject 10	89.1	0.90	0.91	0.94	0.85	0.092
Subject 11	79.6	0.88	0.96	0.93	0.93	0.086
Subject 12	75.7	0.92	0.82	0.94	0.90	0.080
Subject 13	92.5	0.79	0.81	0.97	0.94	0.088
Subject 14	82.5	0.84	0.86	0.92	0.94	0.096
Subject 15	90.9	0.80	0.84	0.81	0.90	0.080
Overall	96.3	0.96	0.82	0.88	0.91	0.1

tested with the other pieces. Using confusion matrix, we calculated accuracy, precision, specificity, recall, f1-score and p-score. All the metrics performed are represented in Tables 3 and 4 on both data sets. We calculate above parameters using below keywords

T_{positive} is when emotion is expected and is present.
F_{positive} is when emotion is expected and is not present.
T_{negative} is when emotion is not expected and is not present.
F_{negative} is when emotion is not expected and is present.

3 emotional states for SEED and 4 emotional states for DEAP data set are classified using two well known classifiers. The results are validated using the accuracy from selected features on each subject of both data sets. The accuracy, precision, recall, F1-score, p-score and specificity of the training set is given by:

$$\text{Acc} = \frac{T_{\text{positive}} + T_{\text{negative}}}{\text{positive} + \text{negative}}$$

$$\text{Prec} = \frac{T_{\text{positive}}}{T_{\text{positive}} + F_{\text{positive}}}$$

$$F\text{1-score} = 2 \times \frac{\text{prec} - \text{recall}}{\text{prec} + \text{recall}}$$

$$\text{Recall} = \frac{T_{\text{positive}}}{T_{\text{positive}} + F_{\text{negative}}}$$

$$\text{Specificity} = \frac{T_{\text{negative}}}{T_{\text{negative}} + F_{\text{positive}}}.$$

4.1.1 *p*-Score

A t test was performed to measure the p-score (probability score). The t test checks whether the selected function will be useful. The t test is used to evaluate whether the selected feature vector differs from the original or actual feature vector. The value of p is determined by calculating the mean difference between the selected objects and the original object vector. Higher p values indicate the quality factor of the selected item.

4.2 Classification performance

Selected features are now ready to categorize the emotions. We have used two modern multi-class classifiers known for this method. The SVM classifications [37] and k-NN [36] work best for results from classes 3 and 4 of SEED and DEAP data sets, respectively. The dimensions of the selected feature vector are 255×36 and 510×48 for the SEED and DEAP data sets, respectively. The results of the two classifiers are shown in Fig. 6, together with the individual networks and the combined characteristics of the three networks, and show that the selected characteristics provide good classification performance.

4.3 Computational cost and model execution time

All the attributes selected using the proposed method and the computational cost [38] of the trained network are calculated using the following equation. Processing time is shortened due to the number of features selected. Using the iMEMD, feature selection, and feature reduction methods proposed

Table 4 Performance evaluation of DEAP data set using SVM classifier

Subject	Accuracy	Precision	Specificity	Recall	F1-score	p-score
Subject 1	69.3	0.66	0.78	0.79	0.78	0.080
Subject 2	68.6	0.90	0.66	0.67	0.74	0.089
Subject 3	76.7	0.71	0.73	0.79	0.77	0.066
Subject 4	67.2	0.72	0.69	0.66	0.83	0.080
Subject 5	81.0	0.83	0.76	0.67	0.69	0.068
Subject 6	81.2	0.80	0.70	0.65	0.84	0.074
Subject 7	66.3	0.93	0.83	0.84	0.70	0.075
Subject 8	72.8	0.88	0.70	0.73	0.67	0.081
Subject 9	65.7	0.82	0.71	0.69	0.69	0.083
Subject 10	77.0	0.79	0.85	0.79	0.78	0.079
Subject 11	71.8	0.65	0.87	0.73	0.75	0.083
Subject 12	75.5	0.86	0.74	0.67	0.72	0.079
Subject 13	71.0	0.66	0.85	0.81	0.75	0.090
Subject 14	77.9	0.72	0.89	0.65	0.90	0.071
Subject 15	71.8	0.85	0.86	0.93	0.74	0.067
Subject 16	74.2	0.69	0.70	0.64	0.65	0.085
Subject 17	70.9	0.81	0.68	0.65	0.79	0.080
Subject 18	77.1	0.66	0.88	0.83	0.79	0.076
Subject 19	67.4	0.65	0.69	0.84	0.82	0.075
Subject 20	80.2	0.93	0.65	0.75	0.90	0.068
Subject 21	75.4	0.82	0.76	0.79	0.71	0.087
Subject 22	67.4	0.82	0.71	0.85	0.74	0.069
Subject 23	77.5	0.67	0.64	0.85	0.81	0.071
Subject 24	78.4	0.75	0.76	0.83	0.73	0.091
Subject 25	73.7	0.67	0.66	0.92	0.65	0.066
Subject 26	78.8	0.91	0.74	0.81	0.87	0.088
Subject 27	74.9	0.82	0.92	0.71	0.65	0.085
Subject 28	79.3	0.85	0.81	0.64	0.78	0.070
Subject 29	67.4	0.87	0.91	0.91	0.64	0.080
Subject 30	74.2	0.92	0.86	0.77	0.93	0.073
Subject 31	65.2	0.86	0.85	0.71	0.77	0.069
Subject 32	68.0	0.77	0.73	0.73	0.78	0.074
Overall	81.1	0.82	0.77	0.81	0.8	0.099

in this work, emotions can be recognized in a short processing time without interrupting the overall classification performance. Table 5 shows the execution time of the trained network with and without feature selection method:

$$C.C = -\frac{1}{F_t} \sum_{s_i=1}^{S} ([F_s \times ln(f)] + [(1 - F_s) \times ln(1 - f)]), \quad (5)$$

where F_t is the training features, F_s is the features of specific emotional state, S represents the number of emotional states, s_i is the state index, f is the selected feature value

## 4.4 Comparison with recent emotion classification methods

The performance of the proposed EEG-based emotion recognition process is verified based on two sets of reference data. Next, we measure the performance of the proposed technique by classifying the selected objects from support vector machine (SVM) and k-NN (k-Nearest Neighbor). The results in Fig. 6 show that the combined and selected features perform better in the model. The higher the precision of the combined and selected features, the higher the quality. So with this method, one can get accurate precision in shorter processing time. Table 6 shows a comparison of our study with previous techniques using EMD and EEG-based techniques for the SEED and DEAP data sets. The work we have proposed is better than other techniques in terms of

Table 5 Training time of the proposed model

Model	Data set 1: SEED			Data set 2:DEAP		
	Channels	FV	Time (s)	Channels	FV	Time (s)
AlexNet	62	41850 × 4096	1534	32	163840 × 4096	1643
ResNet-50	62	41850 × 4096	1621	32	163840 × 4096	1754
Inception V3	62	41850 × 4096	1798	32	163840 × 4096	1903
Combined	62	125550 × 4096	2841	32	491520 × 4096	3643
Combined-iMEMD	62	89678 × 4096	1922	32	351085 × 4096	2042
Combined-iMEMD-FS-FR	12	225 × 36	18.7	8	510 × 48	22.01
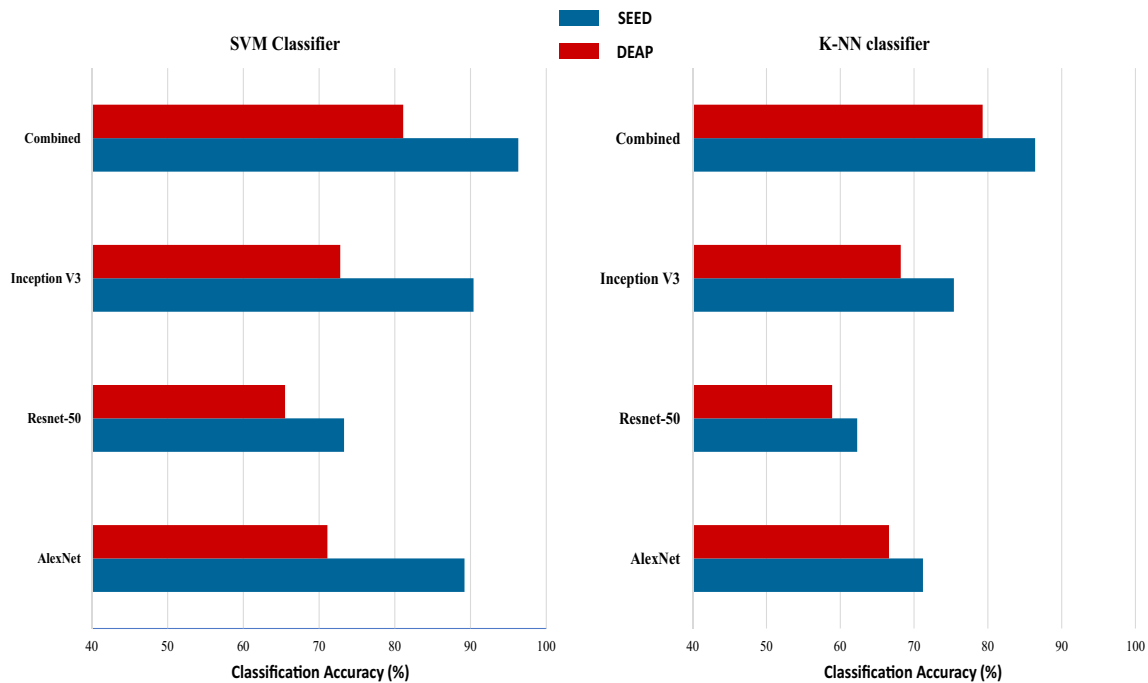
Fig. 6 Average classification accuracy of the proposed method

Table 6 Comparison of the proposed method with recent emotion recognition techniques

Ref.	Model	Features	Channels	Classification accuracy			
				SEED		DEAP	
				SVM	k-NN	SVM	k-NN
[9]	MEMD		10	–	–	60	–
[10]	BEMD		62	71.43	68.52	–	–
[13]	HOS		34	82.1	–	–	–
[14]	EMD	Hilbert Haung transform	62, 32	56	52.6	63.4	57.3
[16]	TEMD	Hilbert Haung transform	62, 32	81.4	79.9	75.3	78
[20]	MEMD	IMF based	86.2	–		75 (ANN)	67
[28]	BoDF	CWT	62, 32	93.8	91.4	77.4	73.6
[15]	MFCC		12, 6	83.5	–	–	71.1
[39]	FEMD	CWT		91.6	83.3	86.4	72.9
Proposed	Multi-model-iMEMD	Spatial	12, 8	96.3	86.4	81.1	79.3
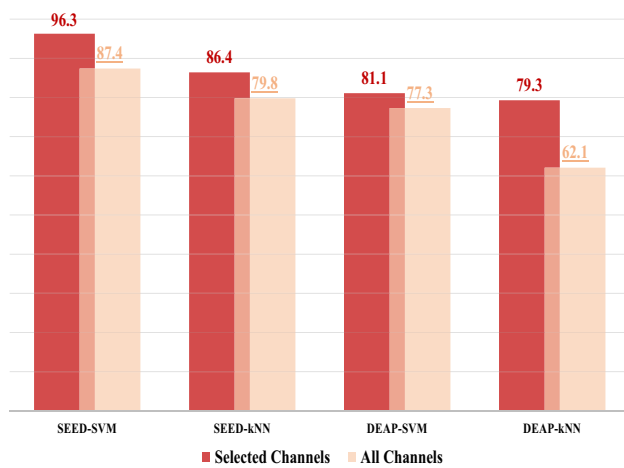
Fig. 7 Comparison of selected vs all channels

classification accuracy and saves computational overhead. The use of three pre-trained networks makes the model simpler and more effective. This table clearly shows that the precision achieved by our model is much better than the previously used model with the same data set and decomposition technique.

The selected channel significantly improves the overall rating performance. In Fig. 7, we can see that the selected channel outperforms all channels. This is because there is an EEG signal, but there are some channels that are not a quality stimulus for emotions. This study shows that channel selection and function help us understand the emotional behavior of the selected channel.

5 Conclusion

This work has proposed an efficient way to select the ability to quickly recognize human emotions. The powerful and intensive functions based on iMEMD are selected from a large table of feature vectors with multiple functions using multiple networks. This model is validated on SEED and DEAP data sets. Reduced feature vectors provide excellent classification performance when verified with SVM and k-NN classifiers. As a result, our feature selection method has been shown to intentionally improve overall classification performance with less computational costs. The proposed model has significantly improved feature extraction properties. Sorting performance has been greatly improved compared to the previous model.

In this study, we propose a method to merge data from multiple neural networks and design a corresponding deep learning model. The proposed technique uses advanced deep neural networks and machine learning methods to select high-quality functions. Channel selection also helps

researchers explore the understanding of emotions on selective channels for accurate recognition of emotions. Future directions of research will explore additional changes to the sentiment analysis framework through the integration of multiple neural networks. The proposed model took less time to test and train the network. For a quick decomposition we use a five step decomposition and this choice is based on the "hit and enforce" method. Therefore, it is better to develop a more systematic way of choosing the level of decomposition.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Jung, Y., Yoon, Y.I.: Multi-level assessment model for wellness service based on human mental stress level. Multimed. Tools Appl. **76**(9), 11305–11317 (2017)
2. Tarnowski, P., Kołodziej, M., Majkowski, A., Rak, R. J.: Combined analysis of GSR and EEG signals for emotion recognition. In 2018 International Interdisciplinary PhD Workshop (IIPhDW) (pp. 137–141). IEEE (2018)
3. Yang, X., Zhang, T., Xu, C., Yan, S., Hossain, M.S., Ghoneim, A.: Deep relative attributes. IEEE Trans. Multimed. **18**(9), 1832–1842 (2016)
4. Faust, O., Hagiwara, Y., Hong, T.J., Lih, O.S., Acharya, U.R.: Deep learning for healthcare applications based on physiological signals: A review. Comput. Methods Programs Biomed. **161**, 1–13 (2018)
5. Hossain, M.S., Muhammad, G., Alamri, A.: Smart healthcare monitoring: a voice pathology detection paradigm for smart cities. Multimed. Syst. **25**(5), 565–575 (2019)
6. Hossain, M.S.: Cloud-supported cyber-physical localization framework for patients monitoring. IEEE Syst. J. **11**(1), 118–127 (2015)
7. Abdulsalam, Y., Hossain, M. S.: COVID-19 networking demand: an auction-based mechanism for automated selection of edge computing services. IEEE Trans. Netw. Sci. Eng. 1–1 (2020). https://doi.org/10.1109/TNSE.2020.3026637
8. Qian, S., Zhang, T., Xu, C., Hossain, M.S.: Social event classification via boosted multimodal supervised latent dirichlet allocation. ACM Trans. Multimed. Comput. Communi. Appl. (TOMM) **11**(2), 1–22 (2015)
9. Kaur, B., Singh, D., Roy, P.P.: EEG based emotion classification mechanism in BCI. Proc. Comput. Sci. **132**, 752–758 (2018)
10. Wei, C., Chen, L.L., Song, Z.Z., Lou, X.G., Li, D.D.: EEG-based emotion recognition using simple recurrent units network and ensemble learning. Biomed. Signal Process. Control **58**, 101756 (2020)
11. Hossain, M.S., Amin, S.U., Alsulaiman, M., Muhammad, G.: Applying deep learning for epilepsy seizure detection and brain

mapping visualization. ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) **15**(1s), 1–17 (2019)

12. Muhammad, G., Hossain, M.S., Kumar, N.: EEG-based pathology detection for home health monitoring. IEEE. J. Sel. Areas Commun. **39**(2), 603–610 (2020)

13. Chakladar, D.D., Chakraborty, S.: EEG based emotion classification using "correlation based subset selection". Biol. Inspired Cognitive Archit. **24**, 98–106 (2018)

14. Eftekhar, A., Toumazou, C., Drakakis, E.M.: Empirical mode decomposition: real-time implementation and applications. J. Signal Process. Syst. **73**(1), 43–58 (2013)

15. Gupta, V., Chopda, M.D., Pachori, R.B.: Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. IEEE Sens. J. **19**(6), 2266–2274 (2018)

16. Ur Rehman, N., Mandic, D.P.: Empirical mode decomposition for trivariate signals. IEEE Trans. Signal Process. **58**(3), 1059–1068 (2009)

17. Tiwari, A., Falk, T.H.: 2019. Fusion of Motif-and spectrum-related features for improved EEG-based emotion recognition. Comput. Intell. Neurosci. (2019)

18. Tong, J., Liu, S., Ke, Y., Gu, B., He, F., Wan, B., Ming, D.: EEG-based emotion recognition using nonlinear feature. In: 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST) (pp. 55–59). IEEE, (2017)

19. Zhuang, N., Zeng, Y., Tong, L., Zhang, C., Zhang, H., Yan, B.: Emotion recognition from EEG signals using multidimensional information in EMD domain. BioMed. Res. Int. **2017**, 9 (2017). https://doi.org/10.1155/2017/8317357

20. Mert, A., Akan, A.: Emotion recognition from EEG signals by using multivariate empirical mode decomposition. Pattern Anal. Appl. **21**(1), 81–89 (2018)

21. Rilling, G., Flandrin, P., Gonçalves, P., Lilly, J.M.: Bivariate empirical mode decomposition. IEEE Signal Process. Lett. **14**(12), 936–939 (2007)

22. Lan, Z., Sourina, O., Wang, L., Scherer, R., Müller-Putz, G.R.: Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets. IEEE Trans. Cognitive Dev. Syst. **11**(1), 85–94 (2018)

23. Wang, Z., Feng, Y., Qi, T., Yang, X., Zhang, J.J.: Adaptive multi-view feature selection for human motion retrieval. Signal Process. **120**, 691–701 (2016)

24. Zhang, Y., Wang, Q., Gong, D.W., Song, X.F.: Nonnegative Laplacian embedding guided subspace learning for unsupervised feature selection. Pattern Recogn. **93**, 337–352 (2019)

25. Chen, Y., Li, H., Hou, L., Wang, J., Bu, X.: An intelligent chatter detection method based on EEMD and feature selection with multi-channel vibration signals. Measurement **127**, 356–365 (2018)

26. Lu, L., Yan, J., de Silva, C.W.: Feature selection for ECG signal processing using improved genetic algorithm and empirical mode decomposition. Measurement **94**, 372–381 (2016)

27. Hossain, M.S., Muhammad, G.: Audio-visual emotion recognition using multi-directional regression and Ridgelet transform. J Multimodel User Interfaces **10**, 325–333 (2016)

28. Asghar, M.A., Khan, M.J., Amin, Y., Rizwan, M., Rahman, M., Badnava, S., Mirjavadi, S.S.: EEG-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach. Sensors **19**(23), 5218 (2019)

29. Zaman, S. M. K., Marma, H. U. M., Liang, X.: Broken rotor bar fault diagnosis for induction motors using power spectral density and complex continuous wavelet transform methods. In: 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE) (pp. 1–4). IEEE (2019)

30. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1026–1034), (2015)

31. Xie, C., Shao, Y., Li, X., He, Y.: Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. Sci. Rep. **5**, 16564 (2015)

32. O'Hara, S., Draper, B. A.: Introduction to the Bag of Features Paradigm for Image Classification and Retrieval. arXiv preprint arXiv:1101.3354, (2011)

33. Zheng, W.L., Lu, B.L.: Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Trans. Auton. Ment. Dev. **7**(3), 162–175 (2015)

34. Tsai, C.F., Hsu, Y.F., Lin, C.Y., Lin, W.Y.: Intrusion detection by machine learning: a review. Expert Syst. Appl. **36**(10), 11994–12000 (2009)

35. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Patras, I.: Deap: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2011)

36. Palaniappan, R., Sundaraj, K., Sundaraj, S.: A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals. BMC Bioinform. **15**(1), 223 (2014)

37. Wichakam, I., Vateekul, P.: An evaluation of feature extraction in EEG-based emotion prediction with support vector machines. In: 2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 106–110). IEEE (2014)

38. Aurelio, Y.S., de Almeida, G.M., de Castro, C.L., Braga, A.P.: Learning from imbalanced data sets with weighted cross-entropy function. Neural Process. Lett. **50**(2), 1937–1949 (2019)

39. Asghar, M. A., Khan, M. J., Amin, Y., Akram, A.: EEG-based Emotion Recognition for Multi Channel Fast Empirical Mode Decomposition using VGG-16. In: 2020 International Conference on Engineering and Emerging Technologies (ICEET) (pp. 1–7). IEEE (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.